

On Polynomial Kernels for Integer Linear Programs: Covering, Packing and Feasibility

Stefan Kratsch^{*†}

We study the existence of polynomial kernels for the problem of deciding feasibility of integer linear programs (ILPs), and for finding good solutions for covering and packing ILPs. Our main results are as follows: First, we show that the ILP FEASIBILITY problem admits no polynomial kernelization when parameterized by both the number of variables and the number of constraints, unless $\text{NP} \subseteq \text{coNP/poly}$. This extends to the restricted cases of bounded variable degree and bounded number of variables per constraint, and to covering and packing ILPs. Second, we give a polynomial kernelization for the COVER ILP problem, asking for a solution to $Ax \geq b$ with $c^T x \leq k$, parameterized by k , when A is row-sparse; this generalizes a known polynomial kernelization for the special case with 0/1-variables and coefficients (d -HITTING SET).

1. Introduction

This work seeks to extend the theoretical understanding of preprocessing and data reduction for Integer Linear Programs (ILPs). Our motivation lies in the fact that ILPs encompass many important problems, and that ILP solvers, especially CPLEX, are known for their preprocessing to simplify (and shrink) input instances before running the main solving routines (see Atamtürk and Savelsbergh [2] for a survey about modern ILP solvers). When it comes to NP-hard problems, then, formally, being able to reduce *every* instance of some problem would give an efficient algorithm for solving it entirely, and prove $\text{P} = \text{NP}$ (cf. [13]). We avoid this issue by studying the question for efficient preprocessing via the notion of *kernelization* from parameterized complexity [7], which relates the performance of the data reduction to one or more problem-specific parameters, like the number n of variables of an ILP.

A kernelization with respect to some parameter n is an efficient algorithm that given an input instance returns an equivalent instance of size depending only on n ; a *polynomial kernelization* guarantees size polynomial in the parameter (see Section 2 for formal definitions). This notion has been successfully applied to a wide range of problems (see Lokshtanov et al. [19] for a recent survey). A breakthrough result by Bodlaender et al. [3] (using [11]) gave a framework for ruling out polynomial kernels for certain problems, assuming $\text{NP} \not\subseteq \text{coNP/poly}$ (else the polynomial hierarchy collapses).¹

^{*}Technical University Berlin, Germany, stefan.kratsch@tu-berlin.de

[†]Supported by the DFG, research project PREMOD, KR 4286/1.

¹All kernelization lower bounds mentioned in this work are modulo this assumption.

ILP feasibility. Let us first discuss the INTEGER LINEAR PROGRAM FEASIBILITY (ILPF) problem: Given a set of m linear (in)equalities in n variables with integer coefficients, decide whether some integer point $x \in \mathbb{Z}^n$ fulfills all of them. A well-known result of Lenstra [18] gives an $\mathcal{O}(\alpha^{n^3} m^c)$ time algorithm for this problem, later improved, e.g., by Kannan [16] to $\mathcal{O}(n^{\mathcal{O}(n)} m^c)$. We can trivially “reduce” to size $N = \mathcal{O}(n^{\mathcal{O}(n)})$ by observing that Kannan’s algorithm solves all larger instances in polynomial time $\mathcal{O}(Nm^c) = N^{\mathcal{O}(1)}$. Can actual reduction rules give smaller kernels, for example with size polynomial in n ?

It is clear that we can store an ILPF instance, for example, (A, b) , with $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$, asking for $x \geq 0$ with $Ax \leq b$, by encoding all $\mathcal{O}(nm)$ coefficients, which takes $\mathcal{O}(nm \log C)$ bits where C is the largest absolute value among coefficients. Let us check what can be said about polynomial kernels with respect to these parameters for ILPF and the r -row-sparse² variant r -ILPF:

If the row-sparseness is unrestricted, then $\text{ILPF}(n+C)$ and $\text{ILPF}(m+C)$ encompass HITTING SET(n) and HITTING SET(m)³, respectively, which admit no polynomial kernels [6, 14]. What about $\text{ILPF}(n+m)$, which is the maximal open case below the trivial parameter $n+m+\log C$ (see Figure 1)?

With bounded row-sparseness r , things turn out differently: For r -ILPF($n+C$) and r -ILPF($m+\log C$) there are polynomial kernels: The former is not hard and we briefly explain it in Appendix D; the latter is trivial since row-sparseness r entails $n \leq r \cdot m$ and hence the above encoding uses $\mathcal{O}(nm \log C) = \mathcal{O}(rm^2 \log C) = (m + \log C)^{\mathcal{O}(1)}$ bits. It was showed previously that r -ILPF(n) admits no polynomial kernel [17], and it can be seen that the proof works also for r -ILPF($n + \log C$).⁴ Again, this leaves parameter $n+m$ open (see Figure 2).

Our contribution for ILPF is the following theorem which, unfortunately, settles both $\text{ILPF}(n+m)$ and r -ILPF($n+m$) negatively (see Section 3). It can be seen that this completes the picture regarding the existence of polynomial kernels for ILPF and r -ILPF for parameterization by any subset of $n, m, \log C$, and C . The same is true for the column-sparse case q -ILPF (see Figure 3), but we omit a detailed discussion since it is quite similar to the row-sparse case.

Theorem 1. *ILPF($n+m$) does not admit a polynomial kernelization or compression, unless $\text{NP} \subseteq \text{coNP/poly}$. This holds also if each constraint has at most three variables and each variable is in at most three constraints.*

It appears that $\text{ILPF}(n+m)$ is the first problem for which we know that a polynomial kernelization fails solely due to the encoding size of large numbers in the input data (and taking into account our proof that no reduction is possible); an additional parameter $\log C$ would trivially give a polynomial kernel. This of course fits into the picture of hardness results for weight(ed) problems, e.g., W[1]-hardness of SMALL SUBSET SUM(k) where the task is to pick a subset of at most k numbers to match some target sum [8] and the kernelization lower bound for SMALL SUBSET SUM($k+d$) where the value of numbers is upper bound by 2^d [6]; however, in both cases the *number* of weights is not bounded in the parameters. Furthermore, there are lower bounds for weighted graph problems (e.g., [15]), but there the used weights have *value* polynomial in the instance size and hence negligible *encoding size*. We also point out two

²Row-sparseness r : at most r variables per constraint; column-sparseness q : each variable occurs in at most q constraints; we use r and q throughout this work.

³HITTING SET: Given a base set U of size n , a set \mathcal{F} of m subsets of U , and an integer k , find a set of k elements of U that intersects each set in \mathcal{F} (if possible). ILP formulation: Is there $(x)_{u \in U}$ with $\sum_{u \in U} x_u \leq k$, and $\sum_{u \in F} x_u \geq 1$ for all $F \in \mathcal{F}$?

⁴The used cross-composition with t input instances creates coefficients of value $\mathcal{O}(t^2)$ with encoding size $\log C = \mathcal{O}(\log t)$ which is permissible for a cross-composition.

contrasting positive results: A randomized polynomial compression of SUBSET SUM(n) [13], and a randomized reduction of KNAPSACK(n) to many instances of size polynomial in n [21] (the number of instances depends on the bit size of the largest weight).

Covering and packing ILPs. Given the overwhelming amount of negative results for ILPF, we turn to the more restricted cases of covering and packing ILPs (cf. [22]) with the hope of identifying some positive cases:

$$\begin{array}{ll}
 \text{(covering ILP:)} \min & c^T x \\
 \text{s.t.} & Ax \geq b \\
 & x \geq 0
 \end{array}
 \qquad
 \begin{array}{ll}
 \text{(packing ILP:)} \max & c^T x \\
 \text{s.t.} & Ax \leq b \\
 & x \geq 0
 \end{array}$$

Here A , b , and c have non-negative integer entries (coefficients). Feasibility for these ILPs is usually trivial (e.g., $x = 0$ is feasible for packing), and the more interesting question is whether there exist feasible solutions x with small (resp. large) value of $c^T x$. Encompassing many well-studied problems from parameterized complexity, we ask whether $c^T x \leq k$ (resp. $c^T x \geq k$), and parameterize by k ; instances are given as (A, b, c, k) . Unsurprisingly, there are a couple of special cases contained in this setting that have been studied before (e.g., with 0/1-variables and coefficients); some of those are W[1]-hard (and are hence unlikely to have polynomial kernels), whereas others have positive results that we could hope to generalize to the more general ILP setting. To capture the different cases, we use the column-sparseness q and row-sparseness r of the matrix A : taking q and r as constants, additional parameters, or unrestricted values defines different problems. Our main result in this part is a polynomial kernelization for r -COVER ILP(k) (see Section 4); the special case of only 0/1 variables and coefficients is known as r -HITTING SET(k) and admits kernels of size $\mathcal{O}(k^r)$ [10, 1]. Our result also uses the Sunflower Lemma (like [10]), but the reduction arguments for sunflowers of linear constraints are of course more involved than for sets.

Theorem 2. *The r -COVER ILP(k) problem admits a reduction to $\mathcal{O}(k^{r^2+r})$ variables and constraints, and a polynomial kernelization.*

Furthermore, we show how to preprocess instances of PACKING ILP($k + q + r$) and COVER ILP($k + q + r$) to equivalent instances with polynomial in kqr many variables and constraints (see Section 5). For r -PACKING ILP($k + q$) this is extended to a polynomial kernelization, generalizing that for the special case of bounded degree INDEPENDENT SET(k). To put these results into context, we provide an overview containing also the inherited hard cases in Tables 1 and 2; for completeness, a brief discussion of these cases can be found in Appendix C.

2. Preliminaries

A *parameterized problem* over some finite alphabet Σ is a language $\mathcal{P} \subseteq \Sigma^* \times \mathbb{N}$. The problem \mathcal{P} is *fixed-parameter tractable* if $(x, k) \in \mathcal{P}$ can be decided in time $f(k) \cdot (|x| + k)^{\mathcal{O}(1)}$, where f is an arbitrary computable function. A *kernelization* for \mathcal{P} is a polynomial-time algorithm that, given input (x, k) , computes an equivalent instance (x', k') with $|x'|, k' \leq h(k)$ where h is some computable function; K is a *polynomial kernelization* if $h(k)$ is polynomially bounded in k . By relaxing the restriction that the created instance (x', k') must be of the same problem and allow the output to be an instance of any language (i.e., any decision problem) we get the notion of (*polynomial*) *compression*. Almost all lower bounds for kernelization apply also for this weaker notion.

Parameterized complexity of Packing ILP(k)			
		row-sparseness r	
		constant	parameter
column-sparseness q	constant	(PK)	(FPT)
	parameter	PK (Theorem 4)	FPT; $n = \mathcal{O}(kqr)$ and $m = \mathcal{O}(kq^2r)$ (Theorem 4)
	unrestricted	W[1]-hard from Independent Set(k)	(W[1]-hard)

Table 1: “PK” stands for polynomial kernel, “no PK” stands for no polynomial kernel unless $\text{NP} \subseteq \text{coNP/poly}$. All normal-font entries are implied by boldface entries.

Parameterized complexity of Cover ILP(k)			
		row-sparseness r	
		constant	parameter
column-sparseness q	constant	(PK)	(FPT)
	parameter	(PK)	(FPT); $n = \mathcal{O}(kqr)$ and $m = \mathcal{O}(kq)$ (Theorem 5)
	unrestricted	PK (Theorem 2)	FPT but no PK (Proposition 4)

Table 2: “PK” stands for polynomial kernel, “no PK” stands for no polynomial kernel unless $\text{NP} \subseteq \text{coNP/poly}$. All normal-font entries are implied by boldface entries.

We also use the concept of an (OR-)cross-composition of Bodlaender et al. [4] which builds on the breakthrough results of Bodlaender et al. [3] and Fortnow and Santhanam [11] for proving lower bounds for kernelization.

Definition 1 ([4]). An equivalence relation \mathcal{R} on Σ^* is called a *polynomial equivalence relation* if the following two conditions hold:

1. There is a polynomial-time algorithm that decides whether two strings belong to the same equivalence class (time polynomial in $|x| + |y|$ for $x, y \in \Sigma^*$).
2. For any finite set $S \subseteq \Sigma^*$ the equivalence relation \mathcal{R} partitions the elements of S into a number of classes that is polynomially bounded in the size of the largest element of S .

Definition 2 ([4]). Let $L \subseteq \Sigma^*$ be a language, let \mathcal{R} be a polynomial equivalence relation on Σ^* , and let $\mathcal{P} \subseteq \Sigma^* \times \mathbb{N}$ be a parameterized problem. An *OR-cross-composition of L into \mathcal{P}* (with respect to \mathcal{R}) is an algorithm that, given t instances $x_1, x_2, \dots, x_t \in \Sigma^*$ of L that are \mathcal{R} -equivalent, takes time polynomial in $\sum_{i=1}^t |x_i|$ and outputs an instance $(y, k) \in \Sigma^* \times \mathbb{N}$ such that:

1. The parameter value k is polynomially bounded in $\max_i |x_i| + \log t$.

2. The instance (y, k) is **yes** if and only if *at least one* instance x_i is **yes**.

We then say that L OR-cross-composes into \mathcal{P} .

Theorem 3 ([4]). *If an NP-hard language L OR-cross-composes into the parameterized problem \mathcal{P} , then \mathcal{P} does not admit a polynomial kernelization or compression unless $\text{NP} \subseteq \text{coNP/poly}$ and the polynomial hierarchy collapses.*

3. A kernel lower bound for ILPs with few coefficients

In this section, we prove Theorem 1, i.e., that INTEGER LINEAR PROGRAM FEASIBILITY($n+m$) admits no polynomial kernelization unless $\text{NP} \subseteq \text{coNP/poly}$. We begin with a technical lemma that expresses multiplication by powers of two in an ILP. The crucial point is that we need multiplication by t different powers of two, but can afford only $\mathcal{O}(\log^c t)$ variables and coefficients (direct products of variables are not legal in linear constraints).

Lemma 1. *Let a, b , and p be variables, and let $b_{\max}, p_{\max} \geq 0$ be integers. Let $\ell = \lceil \log p_{\max} \rceil$. There is a system of $6\ell + 7$ linear constraints with $2\ell - 1$ auxiliary variables such that all integer solutions have $0 \leq b \leq b_{\max}$, $0 \leq p \leq p_{\max}$, and $a = b \cdot 2^p$. Conversely, if these three conditions hold then feasible values for the auxiliary variables exist. The system uses coefficients with bit size $\mathcal{O}(p_{\max} \log b_{\max})$ and all variables have range at most $\{0, \dots, b_{\max} \cdot 2^{p_{\max}}\}$.*

Proof. If $b_{\max} = 0$ then constraints $a = 0$, $b = 0$, $0 \leq p$, and $p \leq p_{\max}$ suffice to prove the lemma; henceforth $b_{\max} \geq 1$. Define a constant $M = b_{\max} \cdot 2^{p_{\max}} + 1$; this will be the coefficient with the largest absolute value (bit size as claimed) and also exceed the value of all involved variables. To enforce this and the claimed bounds on b and p we add the following constraints.

$$\begin{array}{lll} b \geq 0, & p \geq 0, & a \geq 0, \\ b \leq b_{\max}, & p \leq p_{\max}, & a \leq b_{\max} \cdot 2^{p_{\max}} \quad (< M) \end{array}$$

Clearly, we now have $a, b, p < M$. Let $\ell = \lceil \log p_{\max} \rceil$ and add variables $p_0, \dots, p_{\ell-1}$ with range $\{0, 1\}$ (this incurs 2ℓ constraints) together with a constraint

$$p = \sum_{i=0}^{\ell-1} 2^i p_i, \tag{1}$$

which enforces that the p_i form a binary encoding of p . Now, we can rewrite $a = b \cdot 2^p$ as follows (note that this is not added directly as a constraint).

$$a = b \cdot 2^p = b \cdot 2^{\sum_{i=0}^{\ell-1} 2^i p_i} = b \prod_{i=0}^{\ell-1} 2^{2^i p_i} \quad (< M)$$

Our strategy is to enforce the partial products $a_j = b \prod_{i=0}^j 2^{2^i p_i} < M$, for all $j \in \{0, \dots, \ell-1\}$, by enforcing $a_j = 2^{2^j p_j} a_{j-1}$; for notational convenience we identify $a_{\ell-1} := a$ and $a_{-1} := b$ (the remaining $a_0, \dots, a_{\ell-2}$ are new auxiliary variables). We add the following constraints for all $j \in \{0, \dots, \ell-1\}$:

$$a_j \geq a_{j-1} \tag{2}$$

$$a_j \leq 2^{2^j} a_{j-1} \tag{3}$$

$$a_j + M - p_j M \geq 2^{2^j} a_{j-1} \tag{4}$$

$$a_j \leq a_{j-1} + p_j M \tag{5}$$

It suffices to show that $a_j = 2^{2^j p_j} a_{j-1}$ for all $j \in \{0, \dots, \ell-1\}$. First, observe that constraints (2) and (3) restrict a_j to range $\{a_{j-1}, \dots, 2^{2^j} a_{j-1}\}$. Second, consider the effect of $p_j = 0$ respectively $p_j = 1$: If $p_j = 0$ then constraint (4) is trivially fulfilled since $M > b_{\max} \cdot 2^{p_{\max}} \geq 2^{2^j} a_{j-1}$, whereas constraint (5) enforces that $a_j \leq a_{j-1}$; together we get $a_j = a_{j-1} = 2^{2^j p_j} a_{j-1}$ as needed. If instead $p_j = 1$ then constraint (5) is trivially fulfilled whereas constraint (4) enforces $a_j \geq 2^{2^j} a_{j-1}$; together we get $a_j = 2^{2^j} a_{j-1} = 2^{2^j p_j} a_{j-1}$ as needed.

Thus, our set of constraints correctly enforces the intended partial products a_j which implies that $a = a_{\ell-1}$ has to take the desired value $b \cdot 2^p$, as claimed.

For the converse, given the above discussion it is easy to check that setting $p_0, \dots, p_{\ell-1}$ to the binary expansion of p and setting $a_0, \dots, a_{\ell-2}$ to the values of the corresponding partial products gives a feasible assignment. \square

Now we are set up to prove the first part of Theorem 1.

Lemma 2. INTEGER LINEAR PROGRAM FEASIBILITY($n+m$) admits no polynomial kernelization or compression unless $\text{NP} \subseteq \text{coNP/poly}$.

Proof. We give an OR-cross-composition from the NP-hard INDEPENDENT SET problem. The input instances are of the form $(G = (V, E), k)$ where G is a graph and $k \leq |V|$ is an integer, asking whether G contains an independent set of size at least k . For the polynomial equivalence relation \mathcal{R} we let two instances be equivalent if they have the same number of vertices and the same solution size k . It is easy to check that this fulfills Definition 1. For convenience we consider all input graphs to be on vertex set $V = \{1, \dots, n\}$, for some integer n .

Let t \mathcal{R} -equivalent instances $(G_0 = (V, E_0), k), \dots, (G_{t-1} = (V, E_{t-1}), k)$ be given. Without loss of generality we assume that $t = 2^\ell$ for some integer ℓ since otherwise we could copy one instance sufficiently often (at most doubling the input size and not affecting whether at least one of the given instances is **yes**).

Construction. We now describe an instance of INTEGER LINEAR PROGRAM FEASIBILITY that is **yes** if and only if at least one of the instances $(G_0 = (V, E_0), k), \dots, (G_{t-1} = (V, E_{t-1}), k)$ is **yes** for INDEPENDENT SET. We first define an encoding of the t edge sets E_p into $\binom{n}{2}$ integer constants $D(i, j)$, one for each possible edge $\{i, j\} \in \binom{V}{2}$ (throughout the proof we use $1 \leq i < j \leq n$):

$$D(i, j) := \sum_{p=0}^{t-1} 2^p \cdot D(i, j, p), \quad \text{where} \quad D(i, j, p) := \begin{cases} 1 & \text{if } \{i, j\} \in E_p, \\ 0 & \text{else.} \end{cases}$$

In other words, the p -th bit of $D(i, j)$ is one if and only if $\{i, j\}$ is an edge of $G_p = (V, E_p)$. The values $0 \leq D(i, j) \leq 2^t - 1$ will be used as constants in the ILP that we construct next.

1. We start with a single variable p that is intended for choosing an instance number; its range is $\{0, \dots, t-1\}$. The rest of the ILP will be made in such a way that it ensures that a feasible solution for the ILP will imply that instance (G_p, k) is **yes** for INDEPENDENT SET, and vice versa.
2. Now we will add constraints that allow us to extract the necessary information regarding which of the possible edges $\{i, j\}$ are present in graph G_p . Recall that the p -th bit of the constant $D(i, j)$ encodes this. For convenience, we will derive the needed constraints and argue their correctness right away. For each possible edge $\{i, j\}$ with $1 \leq i < j \leq n$ we introduce a variable $e_{i,j}$ with the goal of enforcing $e_{i,j} = 1$ if $\{i, j\} \in E_p$, and $e_{i,j} = 0$ else.

Let i, j with $1 \leq i < j \leq n$ be fixed (we apply the following for all these choices). Clearly,

$$D(i, j) = \sum_{s=0}^{p-1} 2^s D(i, j, s) + 2^p D(i, j, p) + \sum_{s=p+1}^{t-1} 2^s D(i, j, s).$$

We are of course interested in the $2^p D(i, j, p)$ term, which takes value either 0 or 2^p . Since the first term $\sum_{s=0}^{p-1} 2^s D(i, j, s)$ is at most $2^p - 1$ and the last is a multiple of 2^{p+1} , we will extract it via a constraint

$$D(i, j) = \alpha + \beta + \gamma, \tag{6}$$

assuming that we can enforce the following conditions: (i) $0 \leq \alpha \leq 2^p - 1$, (ii) $\beta \in \{0, 2^p\}$, and (iii) $\gamma \in \{0, 2^{p+1}, 2 \cdot 2^{p+1}, \dots\}$. (Note that we use new variables $\alpha_{i,j}, \beta_{i,j}, \gamma_{i,j}, \delta_{i,j}, \varepsilon_{i,j}$ for each choice of i and j , but in the construction the indices are omitted for readability.) This is where Lemma 1 comes into the picture as it permits us to enforce the creation of the required values and range restrictions without using overly many variables and constraints. For (i) we add a variable δ and enforce $\delta = 2^p$ by using Lemma 1 on $a = \delta$, $b = b_{\max} = 1$, and p with $p_{\max} = t - 1$. We then add constraints $\alpha \geq 0$ and $\alpha \leq \delta - 1$. For (ii) we apply the lemma on $a = \beta$, $b = e_{i,j}$ with $b_{\max} = 1$, and p with $p_{\max} = t - 1$, enforcing $\beta = 2^p e_{i,j}$. (Note that we want to get $e_{i,j} = D(i, j, p)$ from $\beta = 2^p \cdot D(i, j, p)$ anyway; this way it is already enforced.) For (iii) we add a new variable $\varepsilon \geq 0$ and apply the lemma on $a = \gamma$, $b = 2\varepsilon$ with $b_{\max} = 2^t - 1$ (formally, this requires a new variable b' and constraint $b' = 2\varepsilon$), and p with $p_{\max} = t - 1$, enforcing $\gamma = 2^p 2\varepsilon = 2^{p+1}\varepsilon$. The fact that there are no further restrictions on ε effectively allows γ to take on any multiple of 2^{p+1} (and no other values). The upper bound b_{\max} on $b = \varepsilon$ for Lemma 1 comes from $D(i, j) \leq 2^t - 1$ (we do not require larger values of γ and ε since $\gamma \leq D(i, j)$ and $\gamma = 2^p \varepsilon$).

The most costly application of Lemma 1 is for (iii) where we have $b_{\max} = 2^t - 1$ and $p_{\max} = t - 1$. This incurs $\mathcal{O}(\log p_{\max}) = \mathcal{O}(\log t)$ additional variables and constraints, and uses coefficient bit size $\mathcal{O}(p_{\max} \log b_{\max}) = \mathcal{O}(t^2)$ (same bounds suffice also for (i) and (ii)). Thus, over all choices of $1 \leq i < j \leq n$, i.e., for getting all the needed edge information, we use $\mathcal{O}(\binom{n}{2} \log t)$ additional variables and constraints. For each variable $e_{i,j}$ our constraints ensure that it is equal one if $\{i, j\}$ is an edge in G_p ; else it has value zero.

3. Now we can finally add the actual edge constraints needed to express the INDEPENDENT SET problem. We add n variables x_1, \dots, x_n with range $\{0, 1\}$, one variable x_i for each vertex $i \in V$. For each possible edge $\{i, j\} \in \binom{V}{2}$, i.e., for all $1 \leq i < j \leq n$ we add the following constraint.

$$x_i + x_j + e_{i,j} \leq 2 \tag{7}$$

Finally, we add a constraint

$$\sum_{i=1}^n x_i \geq k \tag{8}$$

to ensure that we select at least k vertices. This completes the construction.

Let us now check the number of variables and constraints in our created ILP. The number of variables is dominated by the $\mathcal{O}(\binom{n}{2} \log t)$ variables added in Step 2, which come from $\mathcal{O}(\binom{n}{2})$

applications of Lemma 1. The same is true for the number of constraints used. Thus both parameters of our target instance are polynomially bounded in the largest input instance plus $\log t$. Since we postulated no further restrictions on the target ILP (e.g., A and b may have negative coefficients), we will omit a discussion of how to write all constraints as $Ax \leq b$ with $x \geq 0$ (here x is the vector of all variables used) since that is straightforward. The largest bit size of a coefficient is $\mathcal{O}(t^2)$ hence it is easy to see that the whole ILP can be generated in time polynomial in the total input size (which is roughly $\mathcal{O}(t \cdot n^2)$ from the t INDEPENDENT SET instances with n vertices each). It remains to argue correctness of the construction.

Soundness. Assume that the created ILP has at least one feasible solution and pick according values for all variables. The goal is to show that instance number p is **yes** for INDEPENDENT SET. To this end, let $S \subseteq V$ denote the set of vertices i for which $x_i = 1$. By (8) we know that $|S| \geq k$. Now, for any two vertices $i, j \in S$ with $i < j$ it suffices to show that $\{i, j\}$ is not an edge of the graph G_p . By (7) and $i, j \in S$ we know that $e_{i,j} = 0$ since $x_i + x_j + e_{i,j} \leq 2$. As we already showed in the construction that $e_{i,j}$ must equal the p -th bit of $D(i, j)$ this implies that $\{i, j\} \notin E_p$, as needed.

Completeness. Now, assume that for some $p^* \in \{0, \dots, t-1\}$ the instance $(G_{p^*} = (V, E_{p^*}), k)$ is **yes** for INDEPENDENT SET and let $S \subseteq V$ be some independent set of size k in G_{p^*} . Set $p = p^*$ and assign the following values to the other variables to get a feasible solution for the ILP:

- Set the variables x_i to one for all $i \in S$ and to zero for $i \in V \setminus S$. Clearly this fulfills (8).
- Set each edge variable $e_{i,j}$ to one if and only if $\{i, j\}$ is an edge of G_p (and zero else). It is easy to see that together with the x -variables this fulfills all constraints (7) since we only have $x_i = x_j = 1$ if $\{i, j\}$ is not an edge in G_p (by choice of S).
- For each choice of $1 \leq i < j \leq n$, set $\delta_{i,j} = 2^p$ and set

$$\begin{aligned}\alpha_{i,j} &= \sum_{s=0}^{p-1} 2^s D(i, j, s), \\ \beta_{i,j} &= 2^p D(i, j, p), \\ \gamma_{i,j} &= \sum_{s=p+1}^{t-1} 2^s D(i, j, s).\end{aligned}$$

Thus $0 \leq \alpha_{i,j} \leq 2^p - 1 = \delta_{i,j} - 1$ and $\beta = 2^p e_{i,j}$. Also, clearly, our choices for α , β , and γ fulfill (6).

As $\gamma_{i,j}$ is a multiple of 2^{p+1} we can set $\varepsilon_{i,j} = \frac{\gamma_{i,j}}{2^{p+1}}$ to agree with $\gamma_{i,j} = 2^{p+1} \varepsilon_{i,j}$ that we enforced via Lemma 1. Similarly, $\delta_{i,j} = 2^p$ and $\beta_{i,j} = 2^p e_{i,j}$ agree with what we enforced by the lemma. Thus, for all three applications of the lemma (per possible edge $\{i, j\}$) it is guaranteed (by the lemma statement) that feasible values for all auxiliary variables can be found.

It follows that the constructed ILP instance has a feasible solution, as required. This completes the correctness part.

Thus we have an OR-cross-composition from the NP-hard INDEPENDENT SET problem to the INTEGER LINEAR PROGRAM FEASIBILITY($n + m$) problem. By Theorem 3, this implies that INTEGER LINEAR PROGRAM FEASIBILITY($n + m$) has no polynomial kernel or compression unless $\text{NP} \subseteq \text{coNP/poly}$. \square

The second part of Theorem 1 is now an easy corollary. Similarly, we get lower bounds for covering and packing ILP with parameter $n + m$; the proofs are postponed to Appendix A.1 and A.2. Getting the sparseness for Corollary 1 is easier than what was needed for r -ILPF(n) in [17] since both number of constraints and number of variables are bounded in the parameter value.

Corollary 1. *INTEGER LINEAR PROGRAM FEASIBILITY($n + m$) restricted to instances that have at most 3 variables per constraint (row-sparseness) and with each variable occurring in at most 3 constraints (column-sparseness) admits no polynomial kernel or compression unless $\text{NP} \subseteq \text{coNP}/\text{poly}$.*

Corollary 2. *COVER ILP($n + m$) and PACKING ILP($n + m$) do not admit polynomial kernelizations or compressions unless $\text{NP} \subseteq \text{coNP}/\text{poly}$.*

4. Polynomial kernelization for row-sparse Cover ILP(k)

In this section, we prove Theorem 2, by giving a polynomial kernelization for r -COVER ILP(k), generalizing polynomial kernelizations for r -HITTING SET(k) [10, 1]. Our result uses the sunflower lemma (stated below) that can also be used for r -HITTING SET (as in [10]). However, the application is complicated by the fact that the replacement rules for a sunflower of constraints are not as clear-cut as for sets in the hitting set case: Constraints forming a sunflower pairwise overlap on the same variables but with (in general) different coefficients; hence no small replacement is implied. Additionally, we have to bound the number of constraints that have exactly the same set of variables with nonzero coefficients, called *scope*, since the sunflower lemma will only be applied to the set of different scopes. The main work lies in the proof of the following lemma; the polynomial kernelization is given as a corollary.

Lemma 3. *The r -COVER ILP(k) problem admits a polynomial-time preprocessing to an equivalent instance with $\mathcal{O}(k^{r^2+r})$ constraints and variables.*

Before we turn to the proof, we give a lemma that captures some initial reduction arguments including a bound on the number of constraints with the same scope.

Lemma 4. *Given an instance (A, b, c, k) of r -COVER ILP(k) we can in polynomial time reduce to an equivalent instance (A', b', c', k) such that:*

1. *No constraint is satisfied if all variables in its scope are zero, i.e., $b'_i \geq 1$.*
2. *The cost function c'^T is restricted to $1 \leq c'_i \leq k$.*
3. *All feasible solutions with $c'^T x \leq k$ have $x_i \in \{0, \dots, k\}$ for all i .*
4. *There are at most $(k + 1)^d$ constraints for any scope of at most d variables.*

Proof. Recall that we are considering covering constraints of the form $A[i, \cdot]x \geq b[i]$. Constraints that are satisfied when all their variables are zero, are satisfied by all $x \geq 0$ and may be discarded (as we already require $x \geq 0$). If $c_i = 0$ for some variable x_i then we can set x_i to arbitrarily large values and thus satisfy all constraints containing it (here we use monotonicity of covering constraints); we may thus safely delete any such variable and all constraints containing it. Similarly, if $c_i > k$ for some x_i then already when setting $x_i \geq 1$ we cannot get cost $c^T x$ at most k ; thus we delete such variables (but not their constraints). Consequently, we get $1 \leq c_i \leq k$ for all remaining variables x_i . This also limits the possible values with $c^T x \leq k$ to $x_i \in \{0, \dots, k\}$.

Finally, consider any set of d variables with more than $(k+1)^d$ constraints having exactly this scope. It is clear that there are only $(k+1)^d$ possible assignments to the variables that do not violate the maximum cost of k . Each constraint can rule out some of those. It is therefore sufficient to keep only one constraint for each infeasible assignment, as all further constraints are redundant and may be deleted, giving the claimed bound. Note that each constraint has at most r variables in its scope, hence we can perform this reduction in time polynomial in n . \square

We recall sunflowers and the sunflower lemma of Erdős and Rado [9].

Definition 3 ([9]). Let \mathcal{F} denote a family of sets. A *sunflower* in \mathcal{F} of cardinality t and with *core* C is a collection of t sets $\{F_1, \dots, F_t\} \subseteq \mathcal{F}$ such that $F_i \cap F_j = C$ for all $i \neq j$. The sets $F_1 \setminus C, \dots, F_t \setminus C$ are called the *petals* of the sunflower; they are pairwise disjoint. The core C may be empty.

Lemma 5 (Sunflower Lemma [9]). *Let \mathcal{F} denote a family of sets each of size d . If the cardinality of \mathcal{F} is greater than $d! \cdot (t-1)^d$ then \mathcal{F} contains a sunflower of cardinality t , and such a sunflower can be found in polynomial time.*

Proof of Lemma 3. To begin with, we apply Lemma 4 in polynomial time. Afterwards, for each constraint scope with $d \leq r$ variables, there are at most $(k+1)^d = \mathcal{O}(k^r)$ constraints with that scope. Now, we will apply the sunflower lemma to the set of all scopes of size d , for each $d \in \{1, \dots, r\}$. If there are more than $d! \cdot (t-1)^d = \mathcal{O}(t^r)$ sets then we find a sunflower consisting of t sets (scopes). We will show how to remove at least one constraint matching one of the scopes, when t is sufficiently large (and polynomially bounded in k).

If, instead, the number of scopes is at most $d! \cdot (t-1)^d$ (for all d) then we can bound the total number of constraints as follows: We have r choices for $d \in \{1, \dots, r\}$. For each d we have at most $d! \cdot (t-1)^d = \mathcal{O}(t^r)$ constraint scopes. For each scope there are at most $\mathcal{O}(k^r)$ constraints. In total this gives a bound of $r \cdot \mathcal{O}(t^r) \cdot \mathcal{O}(k^r) = \mathcal{O}(t^r \cdot k^r)$. Since each constraint has at most r variables we get the same bound (in \mathcal{O} -notation) for the number of variables.

Removing a constraint. Let us now see how to find a redundant constraint when there are more than $d! \cdot (t-1)^d$ constraint scopes for some $1 \leq d \leq r$. We will also derive an appropriate value for t (at least $t > k$). Consider a t -sunflower in the set of constraint scopes of size d (which we can get in polynomial time from the sunflower lemma). Let its core be denoted by $C = \{x_1, \dots, x_s\}$, with $0 \leq s < d \leq r$, and its pairwise disjoint petals by $\{y_{1,s+1}, \dots, y_{1,d}\}, \dots, \{y_{t,s+1}, \dots, y_{t,d}\}$. (Note that $s < d$ is needed since all sets are different, which requires nonempty petals; else they would all equal the core.) Thus there must be constraints in the ILP matching these scopes. We arbitrarily pick one constraint for each scope:

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + \dots + a_{1,s}x_s + a_{1,s+1}y_{1,s+1} + \dots + a_{1,d}y_{1,d} &\geq b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \dots + a_{2,s}x_s + a_{2,s+1}y_{2,s+1} + \dots + a_{2,d}y_{2,d} &\geq b_2 \\ &\vdots \\ a_{t,1}x_1 + a_{t,2}x_2 + \dots + a_{t,s}x_s + a_{t,s+1}y_{t,s+1} + \dots + a_{t,d}y_{t,d} &\geq b_t \end{aligned}$$

Note that to keep notation simple the indexing of the variables and coefficients is only with respect to the sunflower and makes no assumption about the actual numbering within $Ax \geq b$; all arguments are local to the sunflower.

First, let us note that we may return **no** (or a dummy **no**-instance) if the core is empty and $t > k$: That would give us more than k constraints on disjoint sets of variables that each

require at least one nonzero variable; this is impossible at maximum cost k . In the following, assume that $s \geq 1$.

Since each variable takes values from $\{0, 1, \dots, k\}$ (by Lemma 4), there are at most $(k+1)^s$ possible assignments for the s core variables (in fact there are even less since the sum is at most k). It is clear that assigning zero to all core variables does not lead to a feasible solution since each of the t constraints requires at least one nonzero variable and the petal variables are disjoint. However, unlike for r -HITTING SET, assigning one to a single core variable is not necessarily sufficient to satisfy all constraints, and the value of each variable for a constraint might be quite different. Thus, we may not simply replace the constraints of the sunflower by the restriction of one constraint to the core variables.

To cope with this difficulty we employ a marking strategy. We check all $(k+1)^s$ assignments of choosing a value from $\{0, 1, \dots, k\}$ for each core variable. It is possible that some of the constraints are already satisfied by this core assignment due to the monotone behavior of covering constraints. If more than k constraints need an additional nonzero variable (which would be a petal variable) then clearly this core assignment is infeasible. In this case we arbitrarily mark $k+1$ constraints that are not satisfied by the core assignment alone (i.e., if all their other variables would be zero); these serve to prevent the core assignment from being chosen. If at most k constraints are not yet satisfied then we mark all of them. Clearly, in total we mark at most $(k+1) \cdot (k+1)^s$ constraints.

We will now argue that all unmarked constraints can be deleted. Clearly, deletion can only create false positives, so consider a solution to the instance obtained by deleting any unmarked constraint. If the assignment does not also satisfy the removed constraint (where we take value zero for variables that do not occur after deletion) then, in particular, the same is true for the core assignment made by this assignment. Hence, while marking constraints with respect to this core assignment we must have marked $k+1$ other constraints (that are also not satisfied by the core assignment alone), which are hence not deleted. However, these other constraints cannot all be satisfied with a budget of at most k , a contradiction. Thus, deleting all unmarked constraints is safe.

Therefore, if we have a sunflower with core size s and more than $(k+1) \cdot (k+1)^s$ constraints then our marking procedure allows us to delete at least one constraint. Thus, allowing for core size s up to $r-1$, we set $t = (k+1) \cdot (k+1)^{r-1} + 1 = \mathcal{O}(k^r)$. While we can find t -sunflowers (via the sunflower lemma) we can always delete at least one constraint. Our earlier discussion at the start of the proof now gives a bound of $\mathcal{O}(t^r \cdot k^r) = \mathcal{O}(k^{r^2+r})$ on the number of constraints and variables achieved by this reduction process. This completes the proof. \square

Corollary 3. *The r -COVER ILP(k) problem admits a polynomial compression to size $\mathcal{O}(k^{r^2+2r})$ and a polynomial kernelization.*

Proof (sketch). We know how to reduce to $\mathcal{O}(k^{r^2+r})$ constraints in polynomial time. Since each constraint can equivalently be described by the infeasible assignments that it defines for the variables in its scope, we may encode the instance by replacing each constraint on $d \leq r$ variables by a 0/1-table of dimension $(k+1)^d$: Each 0 entry means that the assignment corresponding to this coordinate is infeasible (e.g., entry at $(2, 3, 5)$ tells whether $x_1 = 2$, $x_2 = 3$, and $x_3 = 5$ is feasible for this constraint, when (x_1, x_2, x_3) is its scope); each 1 stands for a feasible entry. For each constraint this requires $(k+1)^d = \mathcal{O}(k^r)$ bits for the table plus $r \cdot \log(k^{r^2+r}) = \mathcal{O}(\log k)$ bits for encoding the scope (there are at most r variables, each one encoded by a number from 1 to $\mathcal{O}(k^{r^2+r})$). The cost function $c^T x$ requires only the storing of $\mathcal{O}(k^{r^2+r})$ integers with values from 0 to k , taking $\mathcal{O}(\log k)$ bits each. In total this gives the claimed size for the compression.

To get a polynomial kernelization let us first observe that the problem described above is clearly in NP since feasible solutions (certificates) can be verified in polynomial time. Thus, by an argument of Bodlaender et al. [5], we get an encoding as an instance of r -COVER ILP(k) by using the implicit Karp reduction whose existence follows from the fact that r -COVER ILP(k) is complete for NP. The size of this instance is polynomial in $\mathcal{O}(k^{r^2+2r})$, and hence polynomial in k . This completes the polynomial kernelization. \square

5. Further results

In this section we state some results for PACKING ILP(k) and COVER ILP(k) with column-sparseness q and row-sparseness r as additional parameters or constants; the proofs are postponed to Appendix B.1 and B.2.

Theorem 4. *PACKING ILP($k + q + r$) is fixed-parameter tractable and admits a polynomial-time preprocessing to $\mathcal{O}(kqr)$ variables and $\mathcal{O}(kq^2r)$ constraints. For r -PACKING ILP($k + q$) this gives a polynomial kernelization.*

Theorem 5. *COVER ILP($k + q + r$) is fixed-parameter tractable and admits a polynomial-time preprocessing to $\mathcal{O}(kqr)$ variables and $\mathcal{O}(kq)$ constraints. (A polynomial kernelization for r -COVER ILP($k + q$) follows from Corollary 3.)*

6. Conclusion

We have studied different problem settings on integer linear programs. For INTEGER LINEAR PROGRAM FEASIBILITY parameterized by the numbers n of variables and m of constraints we ruled out polynomial kernelizations, assuming $\text{NP} \not\subseteq \text{coNP/poly}$. This still holds when both column- and row-sparseness are at most three. Adding further new and old results, this settles the existence of polynomial kernels for ILPF and r -ILPF for parameterization by any subset of $\{n, m, \log C, C\}$ where C is the maximum absolute value of coefficients.

Regarding covering and packing ILPs, we gave polynomial kernels for r -COVER ILP(k), generalizing r -HITTING SET(k), and for r -PACKING ILP($k + q$). Further results and observations give an almost complete picture regarding q and r except for the question about polynomial kernels for parameter $k + q + r$. We recall that for both problems one can reduce to $n, m \leq (kqr)^{\mathcal{O}(1)}$ but parameterization by n and m only admits no polynomial kernelization (Corollary 2).

References

- [1] Faisal N. Abu-Khzam. A kernelization algorithm for d-hitting set. *J. Comput. Syst. Sci.*, 76(7):524–531, 2010.
- [2] Alper Atamtürk and Martin W. P. Savelsbergh. Integer-programming software systems. *Annals OR*, 140(1):67–124, 2005.
- [3] Hans L. Bodlaender, Rodney G. Downey, Michael R. Fellows, and Danny Hermelin. On problems without polynomial kernels. *J. Comput. Syst. Sci.*, 75(8):423–434, 2009.
- [4] Hans L. Bodlaender, Bart M. P. Jansen, and Stefan Kratsch. Cross-composition: A new technique for kernelization lower bounds. In Schwentick and Dürr [23], pages 165–176.

- [5] Hans L. Bodlaender, Stéphan Thomassé, and Anders Yeo. Kernel bounds for disjoint cycles and disjoint paths. *Theor. Comput. Sci.*, 412(35):4570–4578, 2011.
- [6] Michael Dom, Daniel Lokshtanov, and Saket Saurabh. Incompressibility through colors and ids. In *ICALP (1)*, volume 5555 of *LNCS*, pages 378–389. Springer, 2009.
- [7] Rod G. Downey and M. R. Fellows. *Parameterized Complexity (Monographs in Computer Science)*. Springer, November 1998.
- [8] Rodney G. Downey and Michael R. Fellows. Fixed-parameter tractability and completeness ii: On completeness for $W[1]$. *Theor. Comput. Sci.*, 141(1&2):109–131, 1995.
- [9] Paul Erdős and Richard Rado. Intersection theorems for systems of sets. *J. London Math. Soc.*, 35:85–90, 1960.
- [10] Jörg Flum and Martin Grohe. *Parameterized Complexity Theory (Texts in Theoretical Computer Science. An EATCS Series)*. Springer, March 2006.
- [11] Lance Fortnow and Rahul Santhanam. Infeasibility of instance compression and succinct PCPs for NP. *J. Comput. Syst. Sci.*, 77(1):91–106, 2011.
- [12] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.
- [13] Danny Harnik and Moni Naor. On the compressibility of \mathcal{NP} instances and cryptographic applications. *SIAM J. Comput.*, 39(5):1667–1713, 2010.
- [14] Danny Hermelin, Stefan Kratsch, Karolina Soltys, Magnus Wahlström, and Xi Wu. Hierarchies of inefficient kernelizability. *CoRR*, abs/1110.0976, 2011.
- [15] Bart M. P. Jansen and Hans L. Bodlaender. Vertex cover kernelization revisited: Upper and lower bounds for a refined parameter. In Schwentick and Dürr [23], pages 177–188.
- [16] Ravindran Kannan. Minkowski’s convex body theorem and integer programming. *Mathematics of Operations Research*, 12(3):415–440, 1987.
- [17] Stefan Kratsch. On polynomial kernels for sparse integer linear programs. *CoRR*, abs/1302.3494, 2013. To appear in STACS 2013.
- [18] Hendrik W. Lenstra. Integer programming with a fixed number of variables. *Mathematics of Operations Research*, 8:538–548, 1983.
- [19] Daniel Lokshtanov, Neeldhara Misra, and Saket Saurabh. Kernelization - preprocessing with a guarantee. In *The Multivariate Algorithmic Revolution and Beyond*, volume 7370 of *LNCS*, pages 129–161. Springer, 2012.
- [20] Moni Naor, Leonard J. Schulman, and Aravind Srinivasan. Splitters and near-optimal derandomization. In *FOCS*, pages 182–191, 1995.
- [21] Jesper Nederlof, Erik Jan van Leeuwen, and Ruben van der Zwaan. Reducing a target interval to a few exact queries. In *MFCSS*, volume 7464 of *LNCS*, pages 718–727. Springer, 2012.

- [22] Serge A. Plotkin, David B. Shmoys, and Éva Tardos. Fast approximation algorithms for fractional packing and covering problems. In *FOCS*, pages 495–504, 1991.
- [23] Thomas Schwentick and Christoph Dürr, editors. *28th International Symposium on Theoretical Aspects of Computer Science, STACS 2011*, volume 9 of *LIPICs*, 2011.

A. Omitted proofs from Section 3

A.1. Proof of Corollary 1

Proof (sketch). It suffices to show how an instance of INTEGER LINEAR PROGRAM FEASIBILITY($n + m$) can be modified to fulfill the extra sparseness restrictions, while not blowing up the numbers of variants and constraints to more than $(n + m)^{\mathcal{O}(1)}$ (formally this constitutes a polynomial parameter transformation from the unrestricted to the sparse case). We make no particular assumption about the constraints other than demanding that they are expressible as $\sum_{i=1}^n \alpha_i x_i \triangleleft b$ where the x_i are the variables, α_i 's and b are integer coefficients, and $\triangleleft \in \{\leq, \geq, =\}$. By replacing the sums with partial sums $s_j = \sum_{i=1}^j \alpha_i x_i$ computed via constraints $s_j = s_{j-1} + \alpha_j x_j$ (or, formally, $s_j - s_{j-1} - \alpha_j x_j = 0$), we can reduce the number of variables per constraint to three. Per original constraint this requires $\mathcal{O}(n)$ new variables for partial sums as well as $\mathcal{O}(n)$ constraints, for a total of $\mathcal{O}(nm)$ additional variables and constraints. To reduce also the number of occurrences of each variable, it suffices to make copies of the variables. Conveniently, the variables for the partial sums occur only two times each (intuitively, once for forcing their own value and once for enforcing the next value, resp., for the occurrence of s_n in $s_n \triangleleft b$). Each original variable occurs at most once in each of the m original constraints. When moving to the partial sums, this is replaced by at most one occurrence for computing the partial sum needed for a given constraint; thus it suffices to have m copies $x_i(1), \dots, x_i(m)$ to use in the partial sum constraints. Clearly, this can be achieved via

$$x_i = x_i(1) \quad x_i(1) = x_i(2) \quad \dots \quad x_i(m-1) = x_i(m).$$

This gives two occurrences per variable $x_i(j)$ with the third being in place of x_i in a partial sum constraint. We use an additional $\mathcal{O}(nm)$ variables and $\mathcal{O}(nm)$ further constraints for this part.

It is easy to see that we do not affect feasibility of the ILP since we only make equivalent replacements, respectively introduce copies of variables. In total we use $\mathcal{O}(nm)$ additional variables and constraints, meaning that the final row- and column-sparse ILP has polynomial in $n + m$ many variables and constraints as needed. If we would have a polynomial kernel or compression for this problem, then combined with the present reduction this would constitute a polynomial compression for the INTEGER LINEAR PROGRAM FEASIBILITY($n + m$) problem that we started with, implying $\text{NP} \subseteq \text{coNP}/\text{poly}$. This completes the proof. \square

A.2. Proof of Corollary 2

Proof. We prove the corollary for the case of COVER ILP($n+m$); the case of packing ILPs works along the same lines. To this end, we start from the instance created in the proof of Lemma 2 and turn it into a covering ILP plus a specified bound on the maximum target function value. The main properties that we require is that each variable in the created ILP (for which we ask for feasibility) has an upper bound on its range, e.g., $p \leq t - 1$ and $\beta_{i,j} \leq 2^t$, and that all variables are non-negative. For convenience let us write $B(z)$ for the upper bound on the range of any variable z (irrespective of the actual names used in the proof of Lemma 2).

The instance created in the proof of Lemma 2 asks whether a given set of linear constraints admits a feasible point with integer coordinates. For all involved variables z we have $0 \leq z \leq B(z)$ for some a priori known bound $B(z)$. Let n denote the number of variables, and let z_1, \dots, z_n be an enumeration of the variables. Each constraint can be written as

$$\sum_{i=1}^n \alpha_i z_i \triangleleft b,$$

where $\triangleleft \in \{\leq, \geq, =\}$ and where b and each α_i are integer constants. Obviously, each equality constraint can be split into one \leq and one \geq constraint (with the same coefficients). Subsequently, we may turn each \leq constraint into a \geq constraint by multiplying all its coefficients by -1 .

It remains to handle the fact that we may have negative coefficients. This is where the bounds on the variables come in. For each variable z_i with upper bound $B(z_i)$ we add a constraint

$$z_i + \hat{z}_i \geq B(z_i), \quad (9)$$

where \hat{z}_i is a new variable, constrained to non-negative values by $\hat{z}_i \geq 0$. To ensure that we actually get $z_i + \hat{z}_i = B(z_i)$, we let our target function be

$$\min \sum_{i=1}^n z_i + \sum_{i=1}^n \hat{z}_i. \quad (10)$$

By asking for a feasible solution of cost (target function value) at most $\sum_{i=1}^n B(z_i)$ we ensure that all constraints (9) must be fulfilled with equality (these constraints enforce a lower bound of $\sum_{i=1}^n B(z_i)$ on the cost). Thus, for all i , we have $z_i + \hat{z}_i = B(z_i)$. This permits us to replace all occurrences of z_i with negative coefficient, say $-\gamma$ where $\gamma \geq 1$, by plugging in $z_i = B(z_i) - \hat{z}_i$. We get $-\gamma z_i = -\gamma B(z_i) + \gamma \hat{z}_i$; moving $-\gamma B(z_i)$ into the RHS of the inequality (i.e., adding $\gamma B(z_i)$ to both sides) we have thus reduced the number of negative coefficients by one. Afterwards, should there be a negative RHS coefficient, then we may return a dummy **no** instance. (Note however that we only make equivalent transformations so if there was a feasible point for the initial ILP then this translates naturally to a feasible solution for the obtained covering ILP.)

When all this is done, we have only \geq constraints with non-negative coefficients, and all variables are restricted to non-negative values. Thus we have obtained an equivalent covering ILP with $2n$ variables and at most $n + 2m$ constraints, which transfers the kernelization lower bound from Lemma 2 also to COVER ILP($n + m$). The proof for PACKING ILP($n + m$) is analogous. \square

B. Omitted Proofs from Section 5

B.1. Proof of Theorem 4

We give two lemmas that together imply Theorem 4.

Lemma 6. *Let an instance (A, b, c, k) of PACKING ILP(k) be given and let q and r denote the column- and row-sparseness of A . In polynomial time we can compute an equivalent instance (A', b', c', k) such that*

1. *A' has dimension at most $kq^2r \times kqr$, i.e., the new ILP has at most kq^2r constraints and at most kqr variables.*
2. *$1 \leq c_i \leq k - 1$, i.e., each increase in a variable contributes a value of at least one and at most $k - 1$.*
3. *A' is a submatrix of A , and b' and c' are subvectors of b and c respectively (this means also that column- and row-sparseness of A' are at most the same as for A).*

Proof. Let us begin with a few simple observations: First, we may delete all variables x_i with $c_i = 0$ since we have no increase in value $c^T x$ for setting x_i to a nonzero value but it may make constraints harder to fulfill; it is crucial that all constraints are of the form $A[i, \cdot]x \leq b_i$. Second, we check for variables x_i such that setting $x_i = 1$ already makes some constraint infeasible, i.e., the assignment defined by

$$x_j = \begin{cases} 1 & \text{if } j = i, \\ 0 & \text{else,} \end{cases}$$

is infeasible. We may then delete x_i since it cannot contribute to $c^T x$ (for feasible x). Finally, if $c_i \geq k$ for any x_i then we may answer **yes** since we already know that we can set x_i to one (all others to zero) which gives $c^T x \geq k$. Henceforth $1 \leq c_i \leq k - 1$.

Now, let us see how to get the upper bounds of kq^2r and kqr on the number of constraints and variables. We greedily select variables such that no two appear in a shared constraint; we end up with a set S of variables. If $|S| \geq k$ then assigning one to each variable in S is feasible (since each constraint “sees” only one variable with value one), and the pay off $c_i x_i$ per variable x_i is at least one, so the value of this is at least k . Thus, if $|S| \geq k$ then we may answer **yes**. Otherwise, if $|S| < k$ then every variable not in S must be in a shared constraint with some variable $x_i \in S$. Each such constraint has at most $r - 1$ other variables beyond x_i and each $x_i \in S$ appears in at most q constraints. Thus the total number of variables is at most

$$|S| + (r - 1)q|S| \leq kqr.$$

Since each variable appears in at most q constraints this also limits the number of constraints to at most kq^2r . (If some constraint has $b_i > 0$ but no variables with nonzero coefficients then of course the answer is **no** instead.) Since we only delete variables and constraints from the original instance it is clear that, e.g., A' is a submatrix of A . \square

According to the lemma we can always reduce to at most kq^2r constraints and at most kqr variables. Thus, for PACKING ILP($k + q + r$), i.e., parameterized by k , q , and r , this constitutes a preprocessing to an equivalent instance with a polynomial number of coefficients. This also means that the problem is fixed-parameter tractable by an application of Lenstra’s [18] algorithm for bounded dimension ILPs. When r is constant then this reduction gives a polynomial kernelization, as we show next.

Lemma 7. *r -PACKING ILP($k + q$) admits a polynomial kernelization.*

Proof. To begin, we apply the reduction given by Lemma 6. With the goal of finding a feasible point x with $c^T x \geq k$ it is clear that it suffices to use a maximum range of $\{0, \dots, k\}$ for each variable x_i (if x is feasible then so are all x' with $0 \leq x' \leq x$, and a single position of value k suffices). Furthermore, as for r -COVER ILP(k), any constraint on r variables excludes some of the $|\{0, \dots, k\}|^r$ possible assignments. Thus we can follow the same argumentation as in the proof of Corollary 3 for encoding the feasible points: For each of the kq^2r constraints, we encode in $\mathcal{O}(r \log(kqr)) = \mathcal{O}(\log(kq))$ the names of its r variables, followed by a binary array of bit size $|\{0, \dots, k\}|^r = \mathcal{O}(k^r)$ that encodes for all points in $\{0, \dots, k\}^r$ whether they are feasible (according to this constraint). Finally, we add an encoding of the target function, i.e., a sequence of kqr values from $\{0, \dots, k - 1\}$ which takes $\mathcal{O}(kq \log k)$. This gives a compression to bit size

$$\mathcal{O}(kq^2r \cdot (k^r + \log(kq)) + kq \log k) = \mathcal{O}(k^{r+1}q^2 + kq^2 \log(kq)) = (k + q)^{\mathcal{O}(1)}.$$

Again, it can be observed that the encoded problem is in NP, as a candidate solution x can be efficiently verified (solution value and consistency with all local lists of feasible points). Thus, by a Karp reduction back to r -PACKING ILP($k + q$) we get a polynomial kernelization, as the output size is polynomial in the compression obtained above. \square

B.2. Proof of Theorem 5

Now, we will address Theorem 5. The arguments are analogous to those needed for Theorem 4 so we will focus on the differences.

Lemma 8. *Let an instance (A, b, c, k) of COVER ILP(k) be given and let q and r denote the column- and row-sparseness of A . In polynomial time we can compute either an equivalent instance (A', b', c', k) such that*

1. A' has dimension at most $kq \times kqr$, i.e., the new ILP has at most kq constraints and kqr variables.
2. $1 \leq c \leq k$, i.e., each increase in a variable contributes a cost of at least one and at most k .
3. A' is a submatrix of A , and b' and c' are subvectors of b and c respectively (this means also that column- and row-sparseness of A' is at most the same as for A).

Proof. The reduction to $c \geq 1$ and $b \leq 1$ follows analog to Lemma 4. If $c_i > k$ for any variable x_i then we cannot afford to set it to one (or larger) and may delete it from the ILP. Now, each remaining variable x_i that is set to $x_i \geq 1$ costs at least one and each remaining constraint requires such a variable. Since we have a total budget of k and each variable appears in at most q constraints this limits the number of constraints to at most kq (else we return **no**). The number of variables is at most kqr since each constraint contains at most r variables. Clearly, all required operations take only polynomial time and the produced matrix A' and vectors b' and c' are submatrices respectively subvectors of A , b , and c . \square

Thus the COVER ILP($k + q + r$) problem can be reduced to at most kq constraints and at most kqr variables, i.e., to a polynomial number of coefficients. Regarding polynomial kernels for r -COVER ILP($k + q$) we refer to our more general result for r -COVER ILP(k) by Lemma 3 and Corollary 3. Nevertheless, if the column-sparseness q is indeed small, it is of course possible to obtain a polynomial compression and kernelization via Lemma 8 combined with a direct analog of Lemma 7 (the encoding is entirely the same with slightly better bounds due to fewer constraints and variables; the interpretation of the encoded problem is of course as a covering problem).

C. Basic intractable cases for covering and packing ILPs

In this section we briefly discuss the hardness results that can be observed from various hard problems being captured as packing or covering ILP problems. The first two are immediate; we spell out the ILP for W[1]-hardness of r -PACKING ILP(k) for completeness, but omit a detailed discussion.

Proposition 1. *The r -PACKING ILP(k) problem is hard for W[1] for all $r \geq 2$ by reduction from INDEPENDENT SET(k).*

Proof. Given an instance $(G = (V, E), k)$ for INDEPENDENT SET(k), make an indicator variable x_v for each vertex $v \in V$, and create the following ILP:

$$\begin{aligned} \max \quad & \mathbf{1}^T x \\ \text{s.t.} \quad & x_u + x_v \leq 1 \quad \forall \{u, v\} \in E \\ & x \geq 0 \end{aligned}$$

where $\mathbf{1}$ denotes the all one vector of dimension $|V|$. The created PACKING ILP(k) instance asks for a feasible point x with value $\mathbf{1}^T x$ at least k . \square

Proposition 2. *The COVER ILP(k) problem is hard for $W[2]$ by reduction from HITTING SET(k).*

If we restrict the column-sparseness to any constant $q \geq 4$ then both COVER ILP(k) and PACKING ILP(k) remain $W[1]$ -hard.

Proposition 3. *Let $q \geq 4$. The q -COVER ILP(k) and the q -PACKING ILP(k) problem are $W[1]$ -hard.*

Proof. For both problems we give a reduction from SUBSET SUM(k) which is $W[1]$ -hard due to Downey and Fellows [8] (see also [7]). In SUBSET SUM(k) we are given n non-negative integer values, say, a_1, \dots, a_n , along with a target value t and the question is whether there exists a choice of at most k numbers whose sum is exactly t ; the problem is parameterized by k . To simplify the reduction we ask for a subset with *exactly* k numbers (this can be ensured by adding k times the value 0 for padding any subset matching the target sum to cardinality k). We also assume that $a_i \leq t$ for all i since larger values can be discarded.

We begin with the reduction to q -PACKING ILP(k). We create the following packing ILP with variables x_1, \dots, x_n , asking for a feasible point x with $\mathbf{1}^T x \geq k$.

$$\begin{aligned} \max \quad & \mathbf{1}^T x \\ \text{s.t.} \quad & \sum_{i=1}^n x_i \leq k & x \leq 1 \\ & \sum_{i=1}^n a_i x_i \leq t & \sum_{i=1}^n (t - a_i) x_i \leq (k - 1)t \\ & x \geq 0 \end{aligned}$$

Clearly, each variable appears in exactly 4 constraints, apart from $x \geq 0$. (If variables are defined to be 0/1, or if range constraints $x_i \in \{0, 1\}$ are free, then we get $q \geq 3$.) Note that the following equivalence holds; it uses that feasible solutions with value $\sum_{i=1}^n x_i \geq k$ satisfy also $\sum_{i=1}^n x_i = k$.

$$\begin{aligned} \sum_{i=1}^n (t - a_i) x_i \leq (k - 1)t & \Leftrightarrow t \sum_{i=1}^n x_i - \sum_{i=1}^n a_i x_i \leq (k - 1)t \\ & \Leftrightarrow kt - \sum_{i=1}^n a_i x_i \leq (k - 1)t \\ & \Leftrightarrow \sum_{i=1}^n a_i x_i \geq t \end{aligned}$$

Thus feasible solutions with value at least k satisfy also $\sum_{i=1}^n x_i = k$ and $\sum_{i=1}^n a_i x_i = t$. Since each x_i is either zero or one this translates directly to a choice of k numbers a_i whose sum is exactly t . The reverse holds easily, by setting the x_i according to a choice of k numbers with sum exactly t .

For q -COVER ILP(k) the same style of ILP works, except for a caveat: We do not have the ability to enforce $x \leq 1$. While this could be replaced by

$$\sum_{i \neq j} x_i \geq k - 1$$

for all $j \in \{1, \dots, n\}$, this would violate the desired bound on the column-sparseness. If the problem setting permits to define a range for each variable x_i then of course we can add $x_i \in \{0, 1\}$ instead; this would give column-sparseness $q = 3$. Otherwise, using a derandomized form of color-coding (e.g., using splitters [20]), we could first show W[1]-hardness for Multicolored Subset Sum(k) where the numbers are colored $1, \dots, k$ and the task is to match the target sum by picking one number of each color (the back reduction to Subset Sum(k) is easy by encoding the colors into the numbers). Then we could augment our reduction by demanding that the sum over all indicator variables for numbers of the same color be at least one (giving column-sparseness $q = 4$). This naturally enforces that each variable takes value at most one, since we have exactly k colors. Nevertheless, since this is not the main focus of this work, we will not go into the details of this. For completeness we state the covering ILP with 0/1-variables.

$$\begin{array}{ll} \min & \sum_{i=1}^n x_i \\ \text{s.t.} & \sum_{i=1}^n x_i \geq k & x_i \in \{0, 1\} \\ & \sum_{i=1}^n a_i x_i \geq t & \sum_{i=1}^n (t - a_i) x_i \geq (k - 1)t \end{array}$$

Correctness can be argued analogously to the case of q -PACKING ILP(k). □

Finally, for COVER ILP($k + r$), i.e., parameterized by target solution cost plus row-sparseness, it is straightforward to show fixed-parameter tractability but the problem does not admit a polynomial kernelization by encompassing HITTING SET($k + d$), parameterized by solution size and maximum set size d ; for the latter problem Dom et al. [6, Theorem 6] ruled out polynomial kernels, assuming $\text{NP} \not\subseteq \text{coNP/poly}$.

Proposition 4. *The COVER ILP($k + r$) problem is fixed-parameter tractable but admits no polynomial kernel or compression unless $\text{NP} \subseteq \text{coNP/poly}$.*

Proof. Let us first sketch an $\mathcal{O}(r^k(n + m)^{\mathcal{O}(1)})$ time fpt-algorithm following similar algorithms for, e.g., r -HITTING SET(k) [7]. We may assume that the cost function $c^T x$ has $c_i \geq 1$ for all $i \in \{1, \dots, n\}$ (otherwise such variables x_i can be removed along with their constraints, as setting them to arbitrarily large values satisfies all their constraints but costs zero). The algorithm starts from $x = 0$ and tries to find a minimal feasible solution x with $c^T x \leq k$ by branching. In each recursive call with some partial solution x , it picks an arbitrary infeasible constraint and branches on the choice of one of its at most r variables, and increases the value of that variable by one (in x). This increases the cost $c^T x$ by at least one. In each branch, this is

repeated until either a feasible solution is found, or until we have reached the maximum budget of k . It is easy to prove correctness by using the invariant that each level of the branching tree contains at least one (possibly infeasible) solution x with $x \leq x^*$ (position-wise), where x^* is an arbitrary feasible solution with cost $c^T x$ at most k (if one exists). The runtime is clearly dominated by the $\mathcal{O}(r^k)$ leaves of the search tree (times polynomial factors).

The reduction from HITTING SET($k + d$) to COVER ILP($k + r$) is straightforward: Make an indicator variable for each element of the base set and ask for a solution with sum of indicators at most k . For each set, demand that the sum of indicators for its elements is at least 1; this gives row-sparseness $r = d$. \square

D. A polynomial kernelization for r -ILP Feasibility($n + C$)

Proposition 5. *The r -INTEGER LINEAR PROGRAM FEASIBILITY($n + C$) problem admits a polynomial kernelization.*

Proof. The key fact is that there can be only few different constraints in an instance of r -ILPF($n + C$): Any linear constraint on $d \leq r$ variables, say,

$$\alpha_1 x_{i_1} + \alpha_2 x_{i_2} + \dots + \alpha_d x_{i_d} \blacktriangleleft \beta$$

is characterized by the following aspects.

1. The choice of relation symbol $\blacktriangleleft \in \{\leq, \geq, =\}$.
2. The choice of d variables: $\binom{n}{d} = \mathcal{O}(n^d)$.
3. The values of $\alpha_1, \dots, \alpha_d$ and β . By problem restriction, the absolute values of the coefficients are at most C , hence for each one we have $2C + 1$ choices out of $\{-C, \dots, 0, \dots, C\}$. Thus, there are $(2C + 1)^{d+1}$ choices.

Over all $d \leq r$ this gives a total number of $m = \mathcal{O}(n^r C^{r+1})$ different constraints. Thus, discarding all duplicate constraints (same variables and coefficients) and using the canonical encoding in $\mathcal{O}(nm \log C) = \mathcal{O}(n^{r+1} C^{r+1} \log C)$ bits constitutes a polynomial kernelization with respect to $n + C$. \square

E. A polynomial kernelization for q -ILP Feasibility($m + C$)

Proposition 6. *The q -INTEGER LINEAR PROGRAM FEASIBILITY($m + C$) problem admits a polynomial kernelization.*

Proof. The main observation is that there is only a bounded number of different coefficient patterns for the variables since each variable appears only in a constant number of constraints:

1. Each variable appears in some set of up to q constraints, giving less than $(m+1)^q = \mathcal{O}(m^q)$ choices.
2. In each such constraint, it has a coefficient of absolute value at most C , giving $(2C + 1)^q$ choices.

Now, if two variables, say x_i and x_j have the same pattern, then we delete one of them. Let us check that deletion of any single variable x_j for which we have another variable x_i with the same pattern does no harm: It is clear that a solution for the created ILP can be extended to the original by setting $x_j = 0$. Conversely, let $x_i = a_i$ and $x_j = a_j$ be part of a feasible point for the original ILP. For the created ILP, which does not contain x_j , we create a feasible point by letting $x_i = a_i + a_j$, and letting all other coordinates be the same. It is easy to see that due to having the same pattern, all constraints are still satisfied (this is true for any way of splitting $a_i + a_j$ between x_i and x_j).

In total we keep only one variable for each pattern, giving $\mathcal{O}(m^q C^q)$ variables. This allows an encoding in space $\mathcal{O}(nm \log C) = \mathcal{O}(m^{q+1} C^q \log C)$.

A technical note: We have assumed that there is no global constraint $x \geq 0$ that we would not count as separate constraints $x_i \geq 0$ in the parameter m (by nature of writing ILPs as, e.g., $Ax \leq b$, $x \geq 0$ and taking m to be the number of rows of A). If we have $x \geq 0$ then, in the above argument, letting $x_i = a_i + a_j$ still works (since $a_i, a_j \geq 0$) and, similarly, setting $x_j = 0$ is feasible. \square

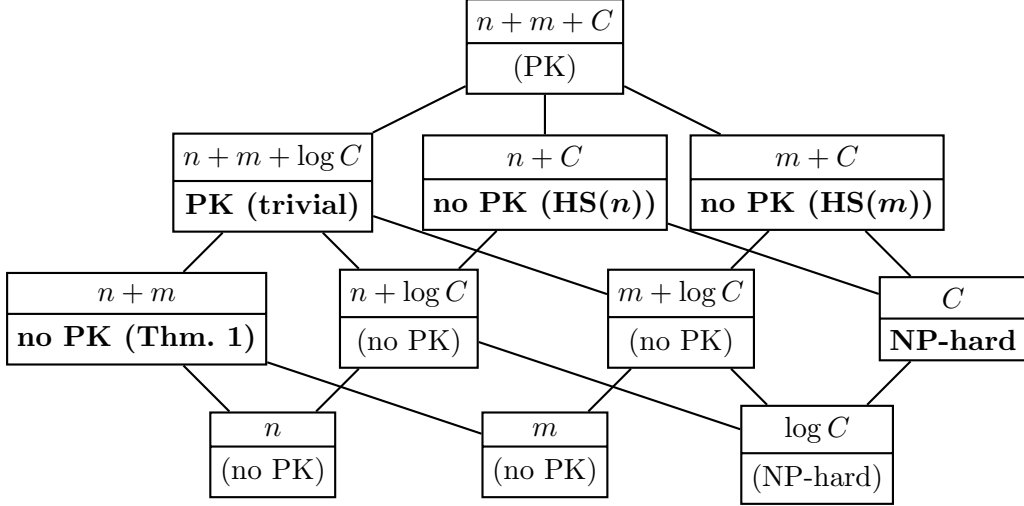


Figure 1: Existence of polynomial kernels for (unrestricted) INTEGER LINEAR PROGRAM FEASIBILITY regarding parameterization by a subset of n , m , $\log C$, and C (depicted in top half of nodes). “PK” indicates a polynomial kernel, “no PK” means that no polynomial kernel is possible unless $\text{NP} \subseteq \text{coNP/poly}$, and “NP-hard” stands for NP-hardness with constant parameter value. Positive results transfer upwards to easier (larger) parameterizations; negative results transfer downwards to harder (smaller) parameterizations. NP-hardness of ILPF(C) for $C = 1$ comes from SATISFIABILITY.

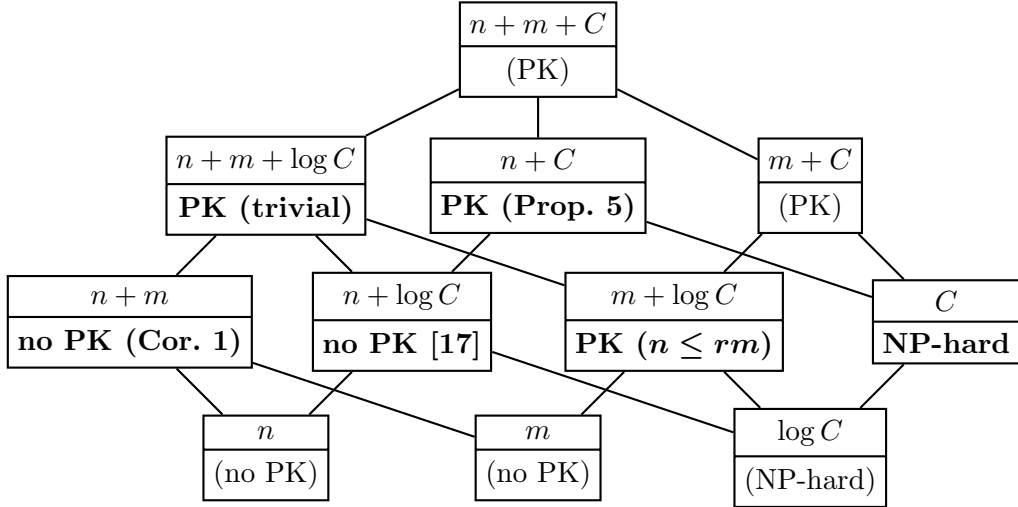


Figure 2: Existence of polynomial kernels for (r -row-sparse) r -INTEGER LINEAR PROGRAM FEASIBILITY regarding parameterization by a subset of n , m , $\log C$, and C (depicted in top half of nodes). “PK” indicates a polynomial kernel, “no PK” means that no polynomial kernel is possible unless $\text{NP} \subseteq \text{coNP/poly}$, and “NP-hard” stands for NP-hardness with constant parameter value. Positive results transfer upwards to easier (larger) parameterizations; negative results transfer downwards to harder (smaller) parameterizations. NP-hardness of r -ILPF(C) for $C = 1$ comes from 3-SATISFIABILITY.

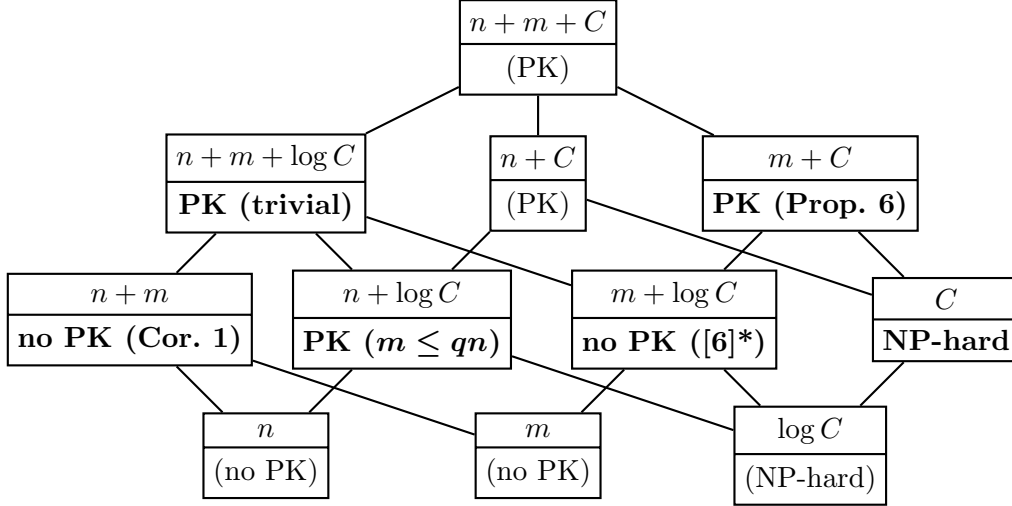


Figure 3: Existence of polynomial kernels for (q -column-sparse) q -INTEGER LINEAR PROGRAM FEASIBILITY regarding parameterization by a subset of n , m , $\log C$, and C (depicted in top half of nodes). “PK” indicates a polynomial kernel, “no PK” means that no polynomial kernel is possible unless $\text{NP} \subseteq \text{coNP/poly}$, and “NP-hard” stands for NP-hardness with constant parameter value. Positive results transfer upwards to easier (larger) parameterizations; negative results transfer downwards to harder (smaller) parameterizations. The lower bound for $q\text{-ILPF}(m + \log C)$ follows from a lower bound for SMALL SUBSET SUM($k + d$) due to Dom et al. [6]: Expressing SUBSET SUM with weights of value at most 2^d requires only $m = \mathcal{O}(1)$ constraints and the encoding size $\log C$ per weight of value at most $C = 2^d$ is $\mathcal{O}(d)$. NP-hardness of $q\text{-ILPF}(C)$ for $C = 1$ comes from SATISFIABILITY restricted to instances with variable degree at most three (cf. [12]).